# What is TEI Conformance and Why Should You Care?

Lou Burnard



## Standards: what's your flavour?

HOW STANDARDS PROLIFERATE: SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)

SITUATION: THERE ARE 14 COMPETING STANDARDS. IH?! RIDICULOUS!
WE NEED TO DEVELOP
ONE UNIVERSAL STANDARD
THAT COVERS EVERYONES
USE CASES. YEAH!



SITUATION: THERE ARE IS COMPETING STANDARDS.

- WKWBFY
- WKWWD
- NUMP



# Standards: what are they good for?

- hermeneutics
- disambiguation
- syntax and semantics



## How to document an encoding scheme

- DTD, RELAX NG, W3C Schema, Schematron
- ... and a lot more besides

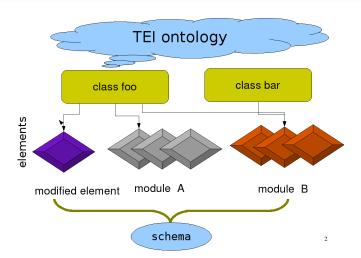


#### A TEI element specification

- identifier(s), glosses, descriptions
- class memberships
- content model and other constrai
- local attributes
- processing model
- usage examples, commentary, references



#### TEI architecture





# The TEI cornucopia





# Dealing with complexity

- selection
- modelling
- revising

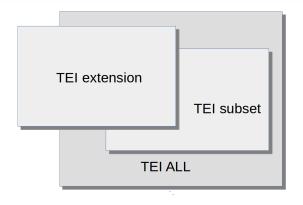
KISS



## Customization is not just selection

- additional constraints
- definitionally true





Varieties of Customization

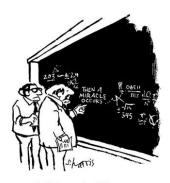


## The Inevitability of Extension

- other vocabularies
- community development
- interchange



#### Unverifiable constraints



"I think you should be more explicit here in step two."

- honesty is the best policy
- explicitness



#### What is TEI conformance?

# The TEI Commandments

- I. Thou shalt have no other encoding scheme but this one
- II. Honour the consensus that thy days may be long in this land
- III. Thou shalt not take the GIs of this scheme in vain
- IV. Thou shalt not commit polysemy

Text Encoding Initiative

650 △ November 1991

1991 version



#### A TEI conformant document (2017 version)

- is a well formed XML document; and
- is valid against one or more schemas, which may be either a TEI subset or a TEI extension; and
- uses elements from the TEI namespace as defined by the TEI Guidelines; and
- fully documents its use of the TEI by means of a TEI-conformant ODD *vel sim*.



## .... why you should care



The challenge has not gone away

- Beyond the DC pidgin
- Rich well-documented markup is algorithmically conformable
- OTA, DTA, SCTA ...

